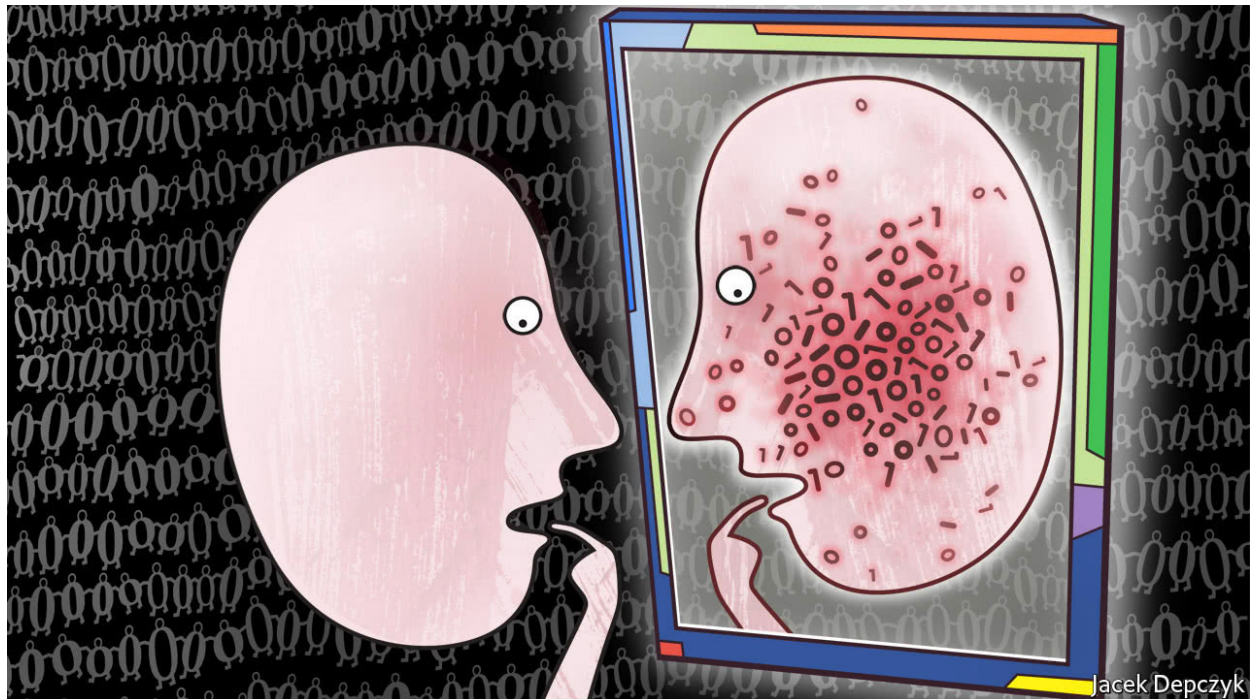


# The digital proletariat Should internet firms pay for the data users currently give away?

[economist.com/news/finance-and-economics/21734390-and-new-paper-proposes-should-data-providers-unionise-should-internet](https://www.economist.com/news/finance-and-economics/21734390-and-new-paper-proposes-should-data-providers-unionise-should-internet)



YOU have multiple jobs, whether you know it or not. Most begin first thing in the morning, when you pick up your phone and begin generating the data that make up Silicon Valley's most important resource. That, at least, is how we ought to think about the role of data-creation in the economy, according to a fascinating new economics paper. We are all digital labourers, helping make possible the fortunes generated by firms like Google and Facebook, the authors argue. If the economy is to function properly in the future—and if a crisis of technological unemployment is to be avoided—we must take account of this, and change the relationship between big internet companies and their users.

Artificial intelligence (AI) is getting better all the time, and stands poised to transform a host of industries, say the authors (Imanol Arrieta Ibarra and Diego Jiménez Hernández, of Stanford University, Leonard Goff, of Columbia University, and Jaron Lanier and Glen Weyl, of Microsoft). But, in order to learn to drive a car or recognise a face, the algorithms that make clever machines tick must usually be trained on massive amounts of data. Internet firms gather these data from users every time they click on a Google search result, say, or issue a command to Alexa. They also Hoover up valuable data from users through the use of tools like reCAPTCHA, which ask visitors to solve problems that are easy for humans but hard for AIs, such as deciphering text from books that machines are unable to parse. That does not just screen out malicious bots, but also helps digitise books. People “pay” for useful free services by providing firms with the data they crave.

These data become part of the firms' capital, and, as such, a fearsome source of competitive advantage. Would-be startups that might challenge internet giants cannot train their AIs without access to the data only those giants possess. Their best hope is often to

be acquired by those very same titans, adding to the problem of uncompetitive markets.

That, for now, AI's contributions to productivity growth are small, the authors say, is partly because of the free-data model, which limits the quality of data gathered. Firms trying to develop useful applications for AI must hope that the data they have are sufficient, or come up with ways to coax users into providing them with better information at no cost. For example, they must pester random people—like those blur-deciphering visitors to websites—into labelling data, and hope that in their annoyance and haste they do not make mistakes.

Even so, as AI improves, the amount of work made vulnerable to displacement by technology grows, and ever more of the value generated in the economy accrues to profitable firms rather than workers. As the authors point out, the share of GDP paid out to workers in wages and salaries—once thought to be relatively stable—has already been declining over the past few decades.

To tackle these problems, they have a radical proposal. Rather than being regarded as capital, data should be treated as labour—and, more specifically, regarded as the property of those who generate such information, unless they agree to provide it to firms in exchange for payment. In such a world, user data might be sold multiple times, to multiple firms, reducing the extent to which data sets serve as barriers to entry. Payments to users for their data would help spread the wealth generated by AI. Firms could also potentially generate better data by paying. Rather than guess what a person is up to as they wander around a shopping centre, for example, firms could ask individuals to share information on which shops were visited and which items were viewed, in exchange for payment. Perhaps most ambitiously, the authors muse that data labour could come to be seen as useful work, conferring the same sort of dignity as paid employment: a desirable side-effect in a possible future of mass automation.

The authors' ideas need fleshing out; their paper, thought-provoking though it is, runs to only five pages. Parts of the envisioned scheme seem impractical. Would people really be interested in taking the time to describe their morning routine or office habits without a substantial monetary inducement (and would their data be valuable enough for firms to pay a substantial amount)? Might not such systems attract data mercenaries, spamming firms with useless junk data simply to make a quick buck?

Nothing to use but your brains

Still, the paper contains essential insights which should frame discussion of data's role in the economy. One concerns the imbalance of power in the market for data. That stems partly from concentration among big internet firms. But it is also because, though data may be extremely valuable in aggregate, an individual's personal data typically are not. For one Facebook user to threaten to deprive Facebook of his data is no threat at all. So effective negotiation with internet firms might require collective action: and the formation, perhaps, of a "data-labour union".

This might have drawbacks. A union might demand too much in compensation for data, for example, impairing the development of useful AIs. It might make all user data freely available and extract compensation by demanding a share of firms' profits; that would rule

out the pay-for-data labour model the authors see as vital to improving data quality. Still, a data union holds potential as a way of solidifying worker power at a time when conventional unions struggle to remain relevant.

Most important, the authors' proposal puts front and centre the collective nature of value in an AI world. Each person becomes something like an oil well, pumping out the fuel that makes the digital economy run. Both fairness and efficiency demand that the distribution of income generated by that fuel should be shared more evenly, according to our contributions. The tricky part is working out how.